Sains Malaysiana 54(10)(2025): 2567-2576
*http://doi.org/10.17576/jsm-2025-5410-18*

# Development of Melioidosis Mapping in Malaysia using Various Relative Risk Models

(Pembangunan Pemetaan Melioidosis di Malaysia menggunakan Pelbagai Model Risiko Relatif)

NAZRINA BINTI AZIZ[1,*], OOI PEI WEN[1], IJLAL BINTI MOHD DIAH[2] & WAQAR HAFEEZ[3]

[1]*School of Quantitative Science, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*
[2]*Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak, Malaysia*
[3]*Haide College, Ocean University of China, Qingdao, Shandong, China*

*Received: 12 March 2025/Accepted: 11 August 2025*

ABSTRACT

Melioidosis is a significant infectious disease caused by *Burkholderia pseudomallei*, which is commonly found in soil and water. The disease is highly endemic in Malaysia, with an estimated 2000 deaths annually, surpassing fatalities from dengue and tuberculosis. Despite its severity, understanding the geographical distribution of melioidosis remains a challenge. In this study, the melioidosis data from 2014 to 2023 in Malaysia were analyzed using Excel and WinBUGS software. Relative risk, a measure comparing the risk in one group to another, was used to map melioidosis risk geographically by using ArcGIS. Four models - Susceptible-Infected-Recovered (SIR), Standardized Morbidity Ratios (SMR), Poisson-Gamma, and Besag-York-Mollie (BYM) - were applied to assess their effectiveness. Mapping highlighted consistently higher relative risk in northern Malaysia, particularly in Perlis and Kedah across multiple models while most other states remained in the very low risk category. Besides, the model performance was compared using the Deviance Information Criterion (DIC) to assess goodness of fit. Findings suggest the Poisson-Gamma model is most suitable and reliable for accurate disease risk mapping to better epidemiological surveillance and targeted public health interventions as it accounts for local variations while maintaining computational efficiency in Malaysia.

Keywords: Disease mapping; epidemiology; melioidosis; relative risk estimation; statistical models

ABSTRAK

Melioidosis ialah penyakit berjangkit yang serius yang disebabkan oleh *Burkholderia pseudomallei*, yang sering ditemui dalam tanah dan air. Penyakit ini adalah endemik di Malaysia dengan anggaran 2000 kematian setiap tahun, melebihi jumlah kematian akibat denggi dan tuberkulosis. Walaupun impaknya besar, pemetaan taburan geografi melioidosis masih menjadi cabaran. Dalam kajian ini, data melioidosis dari tahun 2014 hingga 2023 dianalisis menggunakan perisian Excel dan WinBUGS. Risiko relatif, ukuran membandingkan risiko dalam satu kumpulan dengan kumpulan lain digunakan untuk memetakan risiko melioidosis secara geografi dijana menggunakan ArcGIS. Empat model - Susceptible-Infected-Recovered (SIR), Standardized Morbidity Ratios (SMR), Poisson-Gamma, dan Besag-York-Mollie (BYM) - digunakan untuk menilai ketepatannya. Hasil pemetaan menunjukkan risiko relatif yang lebih tinggi dan tekal di utara Malaysia, khususnya di Perlis dan Kedah merentasi pelbagai model manakala kebanyakan negeri lain kekal dalam kategori risiko sangat rendah. Selain itu, prestasi model dibandingkan dengan Deviance Information Criterion (DIC) untuk menilai kesesuaian. Hasil kajian menunjukkan bahawa model Poisson-Gamma memberikan anggaran risiko relatif yang paling sesuai untuk pemetaan risiko melioidosis dalam membantu meningkatkan pemantauan epidemiologi dan strategi intervensi kesihatan awam yang lebih berkesan kerana ia mengambil kira variasi tempatan sambil mengekalkan kecekapan pengiraan.

Kata kunci: Anggaran risiko relatif; epidemiologi; melioidosis; model statistik; pemetaan penyakit

## INTRODUCTION

Melioidosis, which is also known as Whitmore disease or Nightcliff gardener`s disease, was initially described by Captain Alfred Whitmore in 1911 (Lee 2014). It was caused by *Burkholderia pseudomallei*, a saprophytic Gram-negative bacillus, commonly found in soil and water (Limmathurotsakul & Peacock 2011; Liu, Gee & David 2024; Phillips & Garcia 2024). It is endemic in tropical regions such as Southeast Asia and Northern Australia where it has posed a major public health threat for over 25 years (Limmathurotsakul & Peacock 2011).

In Malaysia, melioidosis was first recorded in 1913 during a laboratory animal outbreak, and since then, it has become endemic with an estimated 2000 annual deaths and

a fatality rate of 48% to 65% among bacteremic patients (Nathan et al. 2018). Melioidosis remains a significant public health threat in Malaysia with high mortality rates, yet spatial risk mapping remains limited. This study aims to fill that gap by evaluating statistical models for effective disease surveillance and intervention planning.

Several epidemiological models are used to estimate disease risk and transmission. These include the Hazard Function, Susceptible-Infected-Recovered (SIR), Standardized Morbidity Ratios (SMR), Poisson-Gamma and Besag-York-Mollie (BYM) models. Hazard Function is the basic computation used to determine the probability that an object will survive to a specific time based on its survival to a precious time (Liberto 2022). The SIR model is used to estimate the disease transmission rate by fitting the models to observed incidence data (Trejo & Hengartner 2022).

The SMR model refers to the ratio of the observed fatalities in a population during a particular time period to the expected deaths over the same period if the study population's age-specific rates were the same as those of the standard population (INED 2024). Poisson-Gamma model is a statistical distribution for overdispersed count data while the BYM model considers the possibility of spatial correlation in the data and the possibility that observations in nearby areas will be more similar than those farther away (Ahlmann-Eltze 2021; Moraga 2019). To study the performance between these four models, Deviance Information Criterion (DIC) is used to estimate candidate models for Bayesian models comparison.

Disease mapping plays a crucial role in visualizing the geographical distribution of disease (Kelley & Breeze 2005). Advancements in digital mapping and geographical information systems (GIS) technology now allow more precise integration of epidemiological data (Koch 2022). In this study, Windows Bayesian Inference Using Gibbs Sampling (WinBUGS) is used for Bayesian modeling via Markov Chain Monte Carlo (MCMC) techniques, while Esri created ArcGIS provides spatial visualization tools to produce risk maps of melioidosis across Malaysia states, enhancing understanding and guiding interventions (Lee 2007; Shaktawat 2020).

## LITERATURE REVIEW

Melioidosis originally identified in 1913 in commemoration of Captain Alfred, who was the first to diagnose the illness, which is also named Whitmore disease or Nightcliff gardener's disease, is endemic in Nightcliff in Australia (Lee 2014; Nathan et al. 2018). In Malaysia, the earliest documented case was recorded in 1911, during an outbreak among laboratory guinea pigs and rabbits at the Institute for Medical Research, in Kuala Lumpur, making it a historical disease for around 111 years (Nathan et al. 2018). The causative agent, *Burkholderia pseudomallei,* is a Gram-negative, saprophytic bacillus commonly found in soil and water (Limmathurotsakul & Peacock 2011; Liu, Gee & David 2024). This organism thrives in tropical and subtropical climates, making Southeast Asia, South Asia, and Northern Australia high risk regions for infection (Lee 2014; Phillips & Garcia 2024). Transmission occurs primarily through contact with contaminated soil or water, especially through torn skin, inhalation or ingestion (Liu, Gee & David 2024). Though rare, human-to-human transmission has been documented (Lee 2014; Liu, Gee & David 2024).

Environmental factors such as heavy rainfall can elevate exposure risk by bringing the bacteria to the surface (Nathan et al. 2018). People with pre-existing conditions such as diabetes, cancer, kidney disease or compromised immunity are particularly vulnerable (Cleveland Clinic 2022). Melioidosis can affect multiple organ systems, leading to serious complications like bone infection (osteomyelitis), joint infection (septic arthritis), collections of pus (abscesses) and acute respiratory distress syndrome (ARDS). The disease may incubate for one to four weeks, although latent infections emerging years later also had the possibility to happen (Cleveland Clinic 2022). Clinical symptoms vary depending on the site of infection and may include fever, cough, skin ulcers, joint pain, confusion, and seizures. Diagnosis involves laboratory tests such as blood cultures, Gram Stain, Polymerase Chain Reaction (PCR) and Enzyme Immunoassay (EIA) (Lee 2014). Treatment requires a two-phase antibiotic regimen: An intensive phase of intravenous antibiotics (Ceftazidime, Meropenem or Imipenem) for a minimum of two weeks, followed by an eradication phase with oral antibiotics (Trimethoprim/Sulfamethoxazole or Amoxicillin/Clavulanic acid) for three months (Cleveland Clinic 2022). Despite effective treatments, mortality rates remain high, particularly among high-risk individuals.

Various epidemiological models are used to understand melioidosis transmission and risk. The Susceptible-Infected-Recovered (SIR) model, introduced by Kermack and McKendrick, divides the population into three compartments: susceptible compartment, $s$ for individuals who have never had a pathogen infection; infected compartment, $i$ for individuals that are currently infected; and removed compartment, $r$ for individuals who have either recovered from the infection and are immune or have passed away are therefore eliminated from the population (Melikechi et al. 2022; Trejo & Hengartner 2022). Given that the SIR model assumes every recovered individuals retains an ongoing immunity to the pathogen, the value of $r$ can be obtained from the summation of $s$, $i$ and $r$ which is equal to 1. The SIR model is easy to compute and interpret, effectively modeling disease progression and outbreak dynamics. However, it assumes equal contact probability among individuals, which may not reflect real-world social structures and contact patterns (Melikechi et al. 2022).

The observed mortality divided by the expected mortality is referred to as standardized morbidity ratio or standardized mortality ratio (SMR) (Litton, Guidet & de Lange 2020). There are two ways to standardized mortality

data which are direct and indirect methods. When the age-sex-specific rates for the study population and the age-sex-structure of the standard population is known, the direct method is used. When the age-specific rates for the study populations are unknown or not available, the indirect method is used. With this model, the ratio of the number of deaths observed in a population during a certain period was compared to the number that would be expected if the study population's age-specific rates were the same as the standard population (Anna 2020; INED 2024). The SMR model is simple and useful for comparing morbidity rates across populations, especially in occupational health, but it lacks flexibility and can be biased if population structures differ (Nicholls 2020).

Poisson-Gamma distribution is a statistical distribution for overdispersed count data which is also known under the name Negative Binomial. The Poisson-Gamma is parametrized by the mean and the overdispersion alpha, unlike the Negative Binomial which is mostly utilised for repeated trials and number of successes or failures. It reduces to the Poisson distribution if alpha is zero (Ahlmann-Eltze 2021). The Poisson-Gamma model allows for covariate adjustments and spatial correlation across neighbouring areas while taking into account the disease's transmission in humans (Jainsankar & Ranjani 2024). Poisson-Gamma model handles overdispersed count data well and allows covariate inclusion, offering more accurate estimates (FasterCapital 2024). Its drawbacks include complexity compared to simple ratios such as the SIR and SMR model, higher computational demands and reliance on the assumption that the Poisson rate follows a Gamma distribution (FasterCapital 2024).

The Besag-York-Mollie (BYM) model is a spatial model that considers the possibility of spatial correlation between data and observations in neighbouring areas that may be more similar than those farther away. This model consists of a spatial random effect that smoothes the data based on a neighbourhood structure, and an unstructured exchangeable component that models uncorrelated noise. Spatio-temporal models that take into consideration not only for spatial structure but also for temporal correlations and spatio-temporal interactions are used in spatio-temporal settings where disease counts are observed over time (Moraga 2019). The BYM model captures spatial patterns and distinguishes structured and unstructured spatial effects, aiding in disease mapping. However, it is statistically complex and requires detailed spatial data (Besag, York & Mollié 1991).

In short, the SIR model is dynamic but sensitive to initial conditions and limiting realism. The SMR model is simple but unstable with sparse data and limited adjustability. The Poisson-Gamma model accounts for overdispersion well but is complex and assumption-dependent. The BYM model incorporates spatial effects effectively but requires advanced methods, spatial data and can be oversmooth.

There are some regional studies that have successfully applied these epidemiological models for disease mapping, particularly in Southeast Asia, Australia, and other tropical regions. For example, Aidalina and Poh Ying (2020) applied a modified SIR model to predict COVID-19 dynamics, demonstrating how movement control measures could flatten the epidemic curve while extending its duration. Maryam, Nor Azah and Zulkifley (2019) applied a Bayesian spatial model incorporating SMR model to analyze lung cancer incidence in Libya, identifying significant spatial clustering. For melioidosis, Limmathurotsakul et al. (2016) used Bayesian spatial modeling - incorporating Poisson-Gamma and hierarchical models - to map disease incidence in Thailand and identify environmental risk factors. Similarly, Currie, Ward and Cheng (2010) conducted a 20-year prospective study in Northern Australia using spatial statistical approaches, including BYM-type models, to evaluate melioidosis clustering and the influence of environmental variables in Darwin.

Deviance Information Criterion (DIC) is a Bayesian method for model comparison, especially after Markov Chain Monte Carlo (MCMC) that WinBUGS can calculate for many models (Li, Yu & Zeng 2020; UC Santa Cruz n.d.). It has been used in a wide range of fields such as biostatistics and ecology (Li, Yu & Zeng 2020). DIC selects a model to minimize a plug-in predictive loss (Li, Yu & Zeng 2020). A lower DIC value is estimated to be the model that would best predict a replicate dataset which has the same structure as that currently observed (UC Santa Cruz n.d.).

Disease mapping is essential for visualizing spatial patterns of melioidosis (Kelley & Breeze 2005). From early hand-drawn maps to modern digital platforms, mapping practices have evolved significantly. The integration of epidemiological data into Geographic Information System (GIS) allows for detailed analysis of disease trends across regions (Koch 2022). Tools like WinBUGS and ArcGIS enhance data visualization and support strategic planning for disease control and prevention.

## MATERIALS AND METHODS

### DATA COLLECTION

To study the relative risk of melioidosis, the melioidosis cases from 2014 until 2023 were used. The melioidosis cases from 2024 were excluded due to the incomplete data records since when the data obtained, it is still in the year 2024 where the year had not yet ended. The melioidosis observed cases, $O_i$ are obtained from the online database, Melioidosis.Info (2024). As the data recorded is taken from different sources and it does not total up all the cases accordingly, thus, the data is transferred and the total observed cases for each state for the year 2014 to 2023 which infected humans only is calculated by using excel software.

From 2014 to 2023, melioidosis cases in Malaysia showed significant regional variation, with the highest peak

in 2017 (777 cases) mostly in Kedah (521 cases) and Perlis (225 cases) due to their extensive agriculture activities (Hassan et al. 2010). Sarawak showed a notable second-highest cases in 2016 (148 cases), likely influenced by its tropical climate and rural landscape (Sia et al. 2021). Urban areas such as W.P Kuala Lumpur had sporadic low cases throughout the period, while states like Kelantan and Johor reported occasional cases. The lowest overall incidence was recorded in 2022 with only 5 cases nationwide, indicating possible fluctuations due to strict movement restrictions and public health measures significantly reduced exposure to environmental sources during COVID-19 pandemic.

The incidence rate is obtained from the paper 'Incidences, Case Fatality Rates and Epidemiology of Melioidosis Worldwide: A Review Paper' where the incidence rate is estimated 5.0 per 100,000 people at risk per year for melioidosis and the population of Malaysia for every state is obtained from OpenDOSM (Department of Statistics Malaysia 2024; Fong et al. 2021). The expected case, $E_i$ is then calculated by using the equation:

$$E_i = \text{Incidence Rate} \times n_i$$

Substitute incidence rate into the formula:

$$E_i = \frac{5.0}{100000} \times n_i$$

where $E_i$ is the expected case of melioidosis for the state $i$ and $n_i$ is the number of population for the state $i$.

### SUSCEPTIBLE-INFECTED-RECOVERED (SIR) MODEL

In this study, the relative risk estimation using the SIR model is calculated manually using excel software. The relative risk calculated is then used in constructing disease mapping using ArcGIS. In SIR model, for $i = 1,2,...,M$ study areas (state in Malaysia) and $j = 1,2,...,T$ periods (year observed), the equations are as follow:

$$S_{i,j} = \mu N_i + (1 - \alpha - \mu)\, S_{i,j-1}$$

$$I_{i,j} = \alpha I_{i,j-1} S_{i,j-1} + (1 - g - \mu - \mu_m)\, I_{i,j-1}$$

$$R_{i,j} = g I_{i,j-1} + (1 - \mu)\, R_{i,j-1}$$

where $S_{i,j}$ is the total number of susceptible persons for area $i$, at time $j$; $I_{i,j}$ is the total number of infectious persons for area $i$, at time $j$ ; $R_{i,j}$ is the total number of recovered persons for area $i$, at time $j$; $g$ is the hazard of an infectious person's being removed (recovery rate); $a$ is the risk of a susceptible person's becoming infective in time period $j$, where $a$ is constant; $\mu$ is the birth and natural death rate of humans per year (assumed equal); and $\mu_m$ is the death rate due to melioidosis per year.

### STANDARDIZED MORBIDITY RATIOS (SMR) MODEL

In this study, the relative risk estimation using the SMR model is calculated manually using excel software. The relative risk calculated is then used in constructing disease mapping using ArcGIS. The relative risk estimation using SMR model in this study is calculated by using formula as follow:

$$\text{SMR} = \frac{O_i}{E_i}$$

where $O_i$ is the number of observed cases and $E_i$ is the number of expected cases.

It should be noted that SMR is a summary measure rather than a full statistical model, and its estimates can be unreliable in areas with small populations. It is best complemented by advanced spatial models.

### POISSON-GAMMA MODEL

In this study, the relative risk estimation using the Poisson-Gamma model is calculated using WinBUGS software. The relative risk calculated is then used in constructing disease mapping using ArcGIS. Theoretically, to calculate Poisson-Gamma model, the steps are as follow:

Let $y_i, i = 1,...,n$ be counts of disease in arbitrary small areas. Also define, for the same areas, expected rates $\{E_i\}$ and relative risks $\{\theta_i\}$. It is assumed that $y_i \sim Pois\,(E_i\, \theta_i)$ given $\theta_i$ are *iid*. Assuming $\theta_i = \theta$, for all $i$ and that the prior distribution of $\theta$, $p(\theta)$ , is $\theta \sim Gamma\,(\alpha,\beta)$ where $E(\theta) = \frac{\alpha}{\beta}$, and $\text{Var}(\theta) = \frac{\alpha}{\beta^2}$. The posterior distribution of $\theta$ is given by,

$$[\theta|y,\, \alpha,\, \beta] = \frac{\beta^*\alpha^*}{\Gamma(\alpha^*)}\, \theta\, \alpha^{*-1} exp(-\theta\beta^*)$$

where $\alpha^* = \sum y_i + \alpha$, $\beta^* = \sum E_i + \beta$. It follows that the predictive distribution is

$$[y^*|y,\, \alpha.\, \beta] = \int f(\, y^*|\theta) f(\theta|\alpha, \beta)\, d\theta$$

$$= \prod_{i=1}^{m} \left[\frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(y_i{}^* + \alpha)}{(E_i + \beta)^{(y_i{}^* + \alpha)}}\right]$$

### BESAG-YORK-MOLLIE (BYM) MODEL

In this study, the relative risk estimation using the BYM model IS calculated using WinBUGS software. The relative risk calculated is then used in constructing disease mapping using ArcGIS. To calculate BYM model, theoretically, the steps are as follow:

$$y_i \sim Poisson\,(e_i\, \theta_i)$$

where $e_i\theta_i$ is the mean of the Poisson distribution; $y_i$ is the observed number of cases in $i$ state; and $e_i$ is the expected number of cases in the $i$ state.

$$log\ \theta_i = \alpha + u_i + v_i$$

where $log\ \theta_i$ is the variability of the log relative risk; $a$ is the overall level of the relative risk; $u_i$ is the spatial random effect, reflecting the correlated heterogeneity; and $v_i$ : random effect, representing the uncorrelated heterogeneity.

The distribution model for the uncorrelated heterogeneity, $v_i$ does not depend on geographic location and is assumed to follow a normal distribution with zero means and a common variance (precision parameter) $\tau_v^2$:

$$v_i \sim N(O,\tau_i^2)$$

For the clustering component, a spatial correlation structure is used, where estimation of the risk in any area depends on neighbouring areas.

$$[u_i|u_j, i \neq j, \tau_u^2] \sim N(\bar{u}_i, \tau_i^2)$$

The mean of the areas bordering area $i$,

$$\bar{u}_i = \frac{1}{\sum_j \omega_{ij}} \sum_j u_j \omega_{ij}$$

$$\tau_i^2 = \frac{\tau_u^2}{\sum_j \omega_{ij}}$$

where $u_i$ is the weighted average of the other $u_j$ , $i \neq j$; $\omega_{i,j}$ is the relationship between the area $i$ and $j$; and $\tau_v^2, \tau_u^2$ is the precision parameters, control the amount of variability of random effects $v$ and $u$.

## DEVIANCE INFORMATION CRITERION

To compare the performance of models in calculating the relative risk, DIC was used and calculated using WinBUGS. Theoretically, the DIC is calculated as follow:

$$DIC = \bar{D} + pD = \hat{D} + 2\,pD$$

where $\bar{D}$ is the posterior mean of the deviance; $pD$ is the effective number of parameters, $pD = \bar{D} - \hat{D}$; $\hat{D}$ is a point estimate of the deviance obtained by substituting in the posterior means $\bar{\theta}$; and $\hat{D}$ is $-2\ log\ (p(y|\bar{\theta}))$.

## RELATIVE RISK CATEGORIZATION

Theoretically, relative risk less than one indicates a decreased risk in the exposed group compared to the unexposed group, suggesting a protective effect. Relative risk equal to one implies no difference in risk between the exposed and unexposed groups to the disease. Relative risk more than one denotes an increased risk in the exposed group, indicating a potential risk factor (Andrade 2015). In this study, the relative risk was separated into 5 categories as shown in Table 1.

## RESULTS AND DISCUSSION

The relative risk estimation for melioidosis from 2014 to 2023 in Malaysia was calculated using excel software for SIR model and SMR model while for Poisson-Gamma model and BYM model, the relative risk was generated using WinBUGS. The trend of relative risk calculated using the SIR model for melioidosis in Malaysia from 2014 to 2023 is shown in Figure 1(a). For this model, melioidosis cases from 2013 were used as the initial input, with only Perak and W.P. Kuala Lumpur recording cases - 1 and 3, respectively. As a result, only these two states showed relative risk across the ten-year period in the SIR model output.

The unusual trend and extreme value produced by the SIR model can be attributed to its sensitivity to small sample sizes and changes in case numbers (Luque-Fernandez 2018). The SIR model tends to exaggerate risk in areas with sudden fluctuations because it lacks adjustments for variability. For example in W.P. Kuala Lumpur, a spike from 3 cases in 2013 to 86 cases in 2014 led to an unusually high relative risk over the ten years. In contrast, Perak's increase from 1 to 2 cases produced a more logical and acceptable risk estimate.

TABLE 1. Relative risk categorization

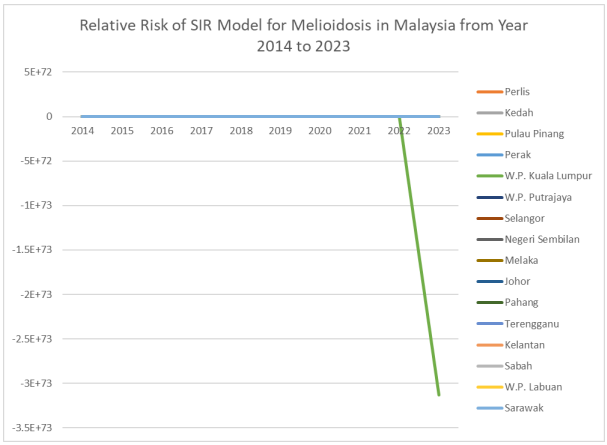
| Category | Relative risk | Colour allocated |
|---|---|---|
| Very low risk | [ 0.0 , 0.5 ) | Light Beige |
| Low risk | [ 0.5 , 1.0 ) | Beige |
| Medium risk | [ 1.0 , 1.5 ) | Light Orange |
| High risk | [ 1.5 , 2.0 ) | Orange Red |
| Very high risk | [ 2.0 , ∞ ) | Dark Red |

Figure 1(b) presents the relative risk trend derived from the SMR model. Since this model considers both observed and expected cases, states with no observed cases displayed no calculated risk. In 2017, Perlis and Kedah recorded the highest risks, at 17.8571 and 4.9619 respectively. The trend generated by the Poisson-Gamma model was shown in Figure 1(c). The trend closely resembles that of the SMR model, with Perlis (16.7800) and Kedah (4.9430) again recording the highest relative risk for melioidosis in 2017. Unlike SMR, the Poisson-Gamma model can still estimate risk even with no observed cases due to its statistical structure.

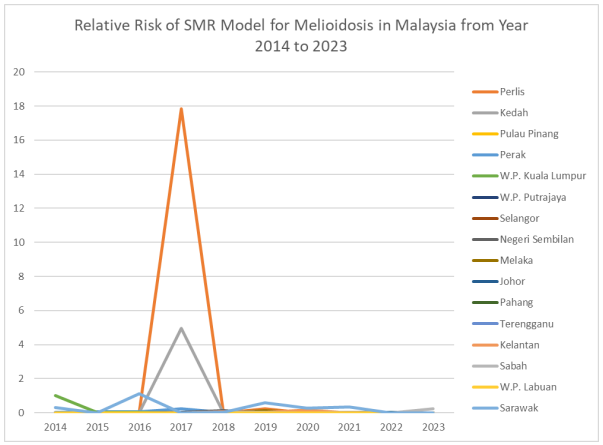
Figure 1(d) shows the relative risk trend using the BYM model. Most states showed a significant downward trend in risk over time. This might result from the over-smoothing characteristics of Bayesian hierarchical models. The BYM model borrows information from neighbouring regions and previous time points, which can suppress high local risk if surrounding areas consistently show low values (Besag, York & Mollié 1991). Additionally, if case reporting was more comprehensive in earlier years and less so in later years, the model might interpret this as a real decline (Lawson 2018). Similar to the Poisson-Gamma model, the BYM model is capable of estimating risk even for states with no observed cases.

Each model serves distinct purposes within the disease analysis workflow. SIR models simulate transmission dynamics, SMR model provides simple risk estimate, Poisson-Gamma model adjuster for overdispersion in count data, and BYM model accounts for spatial dependence. Comparisons should consider these intended functions.

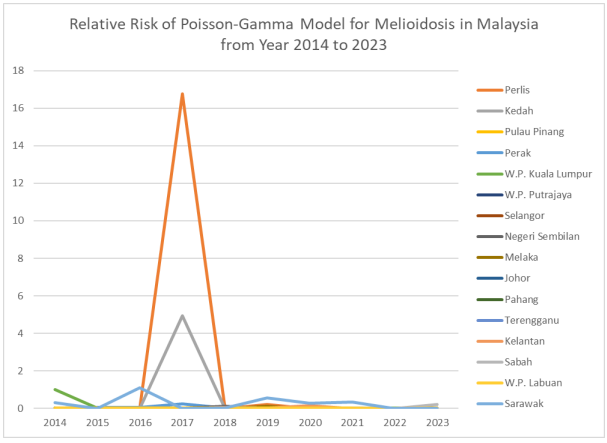
In order to study the performance between the SIR, SMR, Poisson-Gamma and BYM models, the relative risk for melioidosis in 2023 was used to generate disease mapping using ArcGIS. Based on the SIR model results in Table 2, only Perak and W.P. Kuala Lumpur displayed risk - categorized as very high (8.185676) and undefined (-3.1262E+73), respectively. The other 14 states showed no risk, which was still considered as a very low risk category. This was implied in Figure 2(a), where Perak appears in dark red indicating very high risk while the remaining 15 states appear in light beige colour, indicating very low risk.



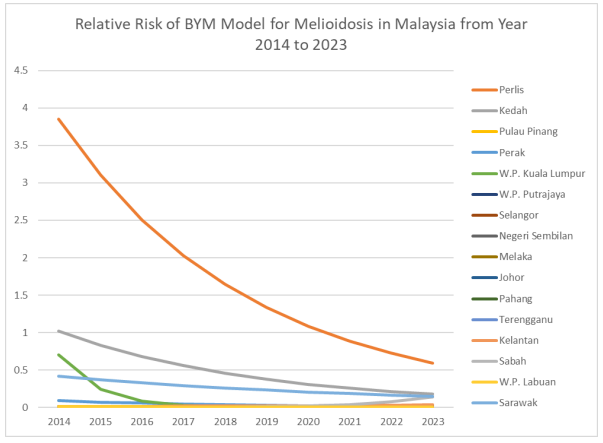

(a) SIR Model




(b) SMR Model




(c) Poisson-Gamma Model




(d) BYM Model

FIGURE 1. Relative risk estimation for melioidosis

Although W.P. Kuala Lumpur had a negative risk value less than 0, it was still grouped under the very low risk category.

From Table 2, the SMR model showed that only Sabah had a measurable risk (0.2167), while the other 15 states showed no risk. All the 16 states were categorized under very low risk with light beige colour displayed in Figure 2(b). For the Poisson-Gamma model, all the 16 states had the relative risk less than 0.5 according to Table 2, categorising them in the very low risk as shown by the light beige colour in Figure 2(c). In the BYM model, Perlis recorded the highest relative risk (0.5969) among all 16 states and was categorized as low risk, represented in beige colour in Figure 2(d) while the remaining 15 states were categorized as very low risk shown in light beige colour.

From Table 2, only Perlis and Perak showed different results between the four models. Perlis was classified as very low risk in SIR (0), SMR (0) and Poisson-Gamma (0.0038) models but was categorized as low risk in the BYM model (0.5969). Conversely, Perak showed a very high risk in SIR model (8.1857) but was classified as very low risk in SMR (0), Poisson-Gamma (0.0005) and BYM (0.0147) models. The SIR model only showed risk in Perak and W.P. Kuala Lumpur as these were the only states with melioidosis cases in 2013. For the SMR model, only Sabah with cases in 2023 showed the risk. In contrast, both the Poisson-Gamma and BYM models produced risk for all the 16 states, including those with no cases as shown in Table 2.

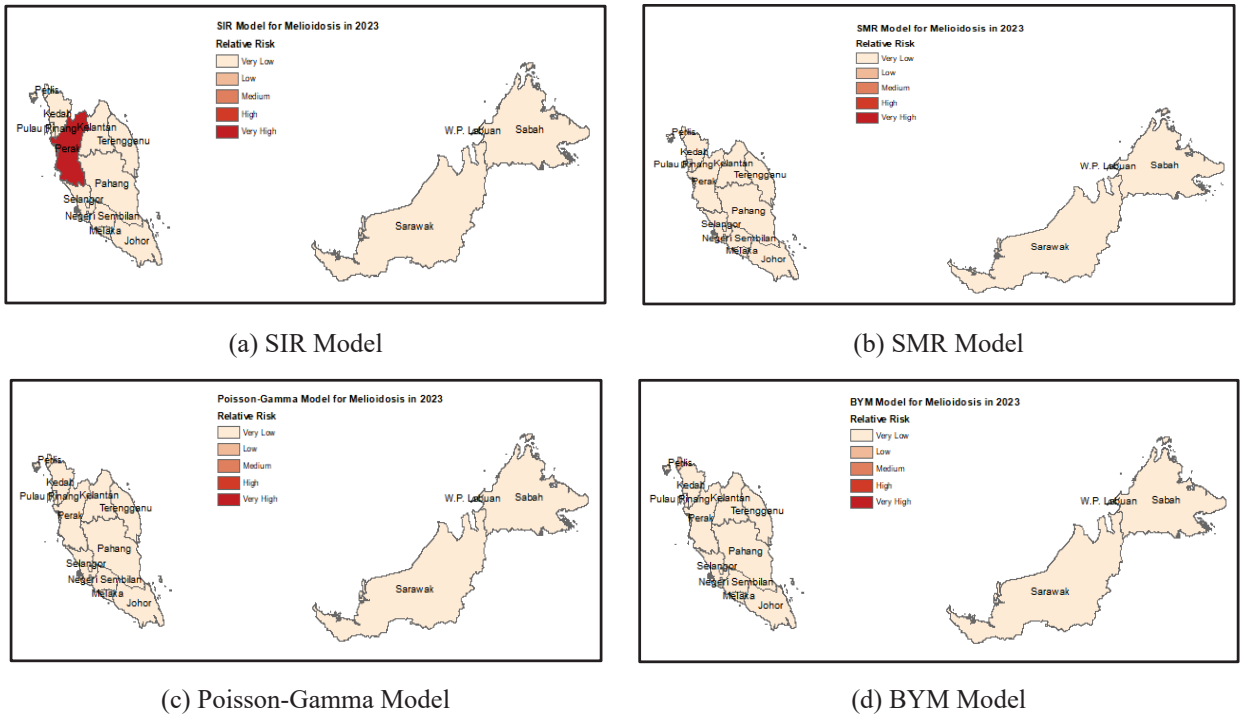
In this study, only Poisson-Gamma model and BYM model allowed for the calculation of DIC using WinBUGS as the SIR and SMR models are not in the family of Bayesian and do not have distribution. Among the four models, the SIR and SMR models were clearly less suitable as they were unable to estimate risk for the state with no observed cases. The SIR model's dependence on initial cases can distort risk estimates when early data are sparse or incomplete. Moreover, the SIR model produced illogical negative relative risk values and unusually high relative risk which further undermine its reliability. Similarly, the SMR model struggles in areas with zero or very low observed cases, making it unstable and less robust in such contexts.

In contrast, the Poisson-Gamma outperformed the BYM model in estimating relative risk for melioidosis. While the BYM model incorporates spatial smoothing by borrowing strength from neighbouring areas, this can obscure true location variation, especially in areas with low case counts. Although the Poisson-Gamma model does not apply spatial smoothing, it provides local data more directly that better reflects the observed data. Most importantly, the Poisson-Gamma model achieved a significantly lower DIC value (262.577) compared to the BYM model (4426.18) as shown in Table 3 was further indicates a better model fit and stronger predictive ability, supporting its suitability for mapping melioidosis risk.

Similar findings have been reported in other melioidosis-endemic regions. Studies from Northern Australia and Thailand demonstrated that Bayesian and overdisperse count models, such as the Poisson-Gamma model, provide more reliable risk estimates than simpler ratio-based methods, particularly in areas with low or

TABLE 2. Relative risk estimation for melioidosis in 2023 in Malaysia

| State | Observed cases | SIR | SMR | Poisson- Gamma | BYM |
|---|---|---|---|---|---|
| Perlis | 0 | 0 | 0 | 0.0038 | 0.5969 |
| Kedah | 0 | 0 | 0 | 0.0006 | 0.1770 |
| Pulau Pinang | 0 | 0 | 0 | 0.0006 | 0.0033 |
| Perak | 0 | 8.1857 | 0 | 0.0005 | 0.0147 |
| W.P. Kuala Lumpur | 0 | -3.1262E+73 | 0 | 0.0006 | 0.0001 |
| W.P. Putrajaya | 0 | 0 | 0 | 0.0109 | 0.0033 |
| Selangor | 0 | 0 | 0 | 0.0002 | 0.0030 |
| Negeri Sembilan | 0 | 0 | 0 | 0.0010 | 0.0099 |
| Melaka | 0 | 0 | 0 | 0.0013 | 0.0154 |
| Johor | 0 | 0 | 0 | 0.0003 | 0.0055 |
| Pahang | 0 | 0 | 0 | 0.0007 | 0.0023 |
| Terengganu | 0 | 0 | 0 | 0.0010 | 0.0081 |
| Kelantan | 0 | 0 | 0 | 0.0007 | 0.0352 |
| Sabah | 39 | 0 | 0.2167 | 0.2164 | 0.1398 |
| W.P. Labuan | 0 | 0 | 0 | 0.0109 | 0.0052 |
| Sarawak | 0 | 0 | 0 | 0.0005 | 0.1505 |




(a) SIR Model

(b) SMR Model

(c) Poisson-Gamma Model

(d) BYM Model

FIGURE 2. Disease mapping for melioidosis in 2023 in Malaysia

TABLE 3. DIC for SIR Model, SMR Model, Poisson-Gamma Model and BYM Model

| | SIR | SMR | Poisson-Gamma | BYM |
|---|---|---|---|---|
| DIC | - | - | 262.577 | 4426.18 |

variable case numbers (Currie, Ward & Cheng 2010; Limmathurotsakul et al. 2016). In Malaysia, prior research has focused more on descriptive mapping without advanced modeling. This study extends local research by confirming that the Poisson-Gamma model offers more accurate and stable risk estimates, especially in states with sparse or no reported cases (Puthucheary 2009). The consistency of high-risk areas identified - such as parts of Perlis and Sabah - also aligns with known melioidosis distributions in Malaysia, further validating the model's performance.

## CONCLUSION

Melioidosis is an infectious disease caused by *Burkholderia pseudomallei* which can be found in soil and water (Liu, Gee & David 2024). It can be transmitted through natural infection which is direct contact of damaged skin with contaminated soil and water, ingestion, and inhalation. Based on the melioidosis' incidence and mortality, it is predicted that about 2000 persons die of melioidosis annually in Malaysia, which is significantly greater than the number of deaths from dengue or tuberculosis infections (Nathan et al. 2018).

In epidemiology and biostatistics, there are various models and ratios that can be used to understand and quantify the risk associated with diseases. In this study, the Susceptible-Infected-Recovered (SIR) model, Standardized Morbidity Ratios (SMR) model, Poisson-Gamma model and Besag-York-Mollie (BYM) model was chosen. Excel software was used to calculate the relative risk of melioidosis for SIR model and SMR model while WinBUGS was selected as the tool to generate the relative risk of melioidosis for Poisson-Gamma model and BYM model. The four models had given different results in terms of the relative risk estimation. The SIR model gave an unusual value of relative risk, both SIR and SMR models only could generate the relative risk for the states with melioidosis cases, both Poisson-Gamma and BYM models can generate the relative risk for all the 16 states but BYM model showed a downward trend of relative risk for melioidosis from 2014 to 2023.

After obtaining the relative risk through the four models, DIC was used to compare the relative risk estimation performance of the four models. Among the four models, the DIC for SIR model and SMR model could

not be calculated due to the lack of distribution while the Poisson-Gamma model was proven to be the best model to calculate the relative risk estimation for melioidosis in this study with its DIC value smaller than BYM model. Besides, ArcGIS was chosen to show visually the melioidosis among the 16 states in Malaysia to construct the melioidosis disease mapping based on the calculated relative risk estimation. In the mapping, the relative risk estimation was classified into 5 categories which were very low risk, low risk, medium risk, high risk, and very high risk. This study demonstrates that the Poisson-Gamma model effectively captures spatial variation in melioidosis risk, making it a valuable tool for mapping and guiding targeted public health action.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to Universiti Utara Malaysia for its continuous support throughout this research. We also extend our appreciation to the Ministry of Health, the Department of Statistics Malaysia, and the World Health Organization for their valuable contributions and data support, which were instrumental in the successful completion of this study. This research was supported by the Fundamental Research Grant Scheme (FRGS/1/2018/STG06/UUM/02/6) under the Ministry of Higher Education (MoHE), Malaysia.

REFERENCES

Ahlmann-Eltze, C. 2021. *Gamma-Poisson Distribution*. https://const-ae.name/post/2021-01-24-gamma-poisson-distribution/gamma-poisson-reference/

Aidalina Mahmud & Poh Ying Lim. 2020. Applying the SEIR Model in forecasting: The COVID-19 trend in Malaysia: A preliminary study. *medRxiv.* https://www.medrxiv.org/content/10.1101/2020.04.14.20065607v1

Andrade, C. 2015. Understanding relative risk, odds ratio, and related terms: As simple as it can get. *J. Clin. Psychiatry* 76(7): e857-861.

Anna, N. 2020. The standardised mortality ratio and how to calculate it. *Students 4 Best Evidence.* https://s4be.cochrane.org/blog/2020/08/26/the-standardised-mortality-ratio-and-how-to-calculate-it/

Besag, J., York, J. & Mollié, A. 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43: 1-20.

Cleveland Clinic. 2022. *Melioidosis*. Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/24051-melioidosis

Currie, B.J., Ward, L. & Cheng, A.C. 2010. The epidemiology and clinical spectrum of melioidosis: 540 cases from the 20 year Darwin prospective study. *PLoS Neglected Tropical Diseases* 4(11): e900. https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0000900

Department of Statistics Malaysia. 2024. *The Population of Malaysia*. OpenDOSM. https://open.dosm.gov.my/dashboard/population

FasterCapital. 2024. Poisson Distribution: How to Use the Poisson Distribution to Model the Probability Distribution of Investment Estimation. https://fastercapital.com/content/Poisson-Distribution--How-to-Use-the-Poisson-Distribution-to-Model-the-Probability-Distribution-of-Investment-Estimation.html#Advantages-and-Disadvantages-of-Using-the-Poisson-Distribution

Fong, J.H., Pillai, N., Yap, C.G. & Jahan, N.K. 2021. Incidences, case fatality rates and epidemiology of melioidosis worldwide: A review paper. *Open Access Library Journal* 8: e7537. https://www.scirp.org/journal/paperinformation?paperid=110072

Hassan, M.R., Pani, S.P., Peng, N.P., Voralu, K., Vijayalakshmi, N., Mehanderkar, R., Aziz, N.A. & Michael, E. 2010. Incidence, risk factors and clinical epidemiology of melioidosis: A complex socio-ecological emerging infectious disease in the Alor Setar region of Kedah, Malaysia. *BMC Infect Dis.* 10: 302. doi: 10.1186/1471-2334-10-302

Institut National D'Etudes Demographiques (INED). 2024. *Standardized Mortality Rate*. https://www.ined.fr/en/glossary/standardized-mortality-rate/

Jainsankar, R. & Ranjani, M. 2024. Spatial disease mapping using the Poisson-Gamma model. *Journal of Future Sustainability* 4: 101-106. doi: 10.5267/j.jfs.2024.5.004

Kelley, L.C. & Breeze, R.G. 2005. Investigation of Suspicious Disease Outbreaks. In *Microbial Forensics.* pp. 187-212. Academic Press. https://www.sciencedirect.com/topics/medicine-and-dentistry/disease-mapping

Koch, T. 2022. Disease mapping and innovation: A history from wood-block prints to Web 3.0. *Patterns* 3(6): 100507. doi: 10.1016/j.patter.2022.100507

Lawson, A.B. 2018. *Bayesian Disease Mapping: Hierarchical Modeling in Spatial Epidemiology*. 3rd ed. Boca Raton: CRC Press.

Lee, P.C. 2014. *Melioidosis.* Myhealth Ministry of Health Malaysia. http://www.myhealth.gov.my/en/melioidosis-2/

Lee, S.Y. 2007. Bayesian analysis of mixtures structural equation models with missing data. *Handbook of Latent Variable and Related Models.* North-Holland. pp. 87-107.

Li, Y., Yu, J. & Zeng, T. 2020. Deviance information criterion for latent variable models and misspecified models. *Journal of Econometrics* 216(2): 450-493. https://doi.org/10.1016/j.jeconom.2019.11.002

Liberto, D. 2022. Hazard rate: Definition, how to calculate, and example. *Investopedia*. https://www.investopedia.com/terms/h/hazard-rate.asp

Limmathurotsakul, D. & Peacock, S.J. 2011. Melioidosis: A clinical overview. *British Medical Bulletin* 99: 125-139. doi:10.1093/bmb/ldr007

Limmathurotsakul, D., Golding, N., Dance, D.A., Messina, J.P., Pigott, D.M., Moyes, C.L., Rolim, D.B., Bertherat, E., Day, N.P., Peacock, S.J. & Hay, S.I. 2016. Predicted global distribution of *Burkholderia pseudomallei* and burden of melioidosis. *Nature Microbiology* 1(1): 15008. doi: 10.1038/nmicrobiol.2015.8

Litton, E., Guidet, B. & de Lange, D. 2020. National registries: Lessons learnt from quality improvement initiatives in intensive care. *J. Crit. Care* 60: 311-318. https://www.sciencedirect.com/topics/medicine-and-dentistry/standardized-mortality-ratio

Liu, L., Gee, J. & David, B. 2024. *Melioidosis.* Centers for Disease Control and Prevention. https://wwwnc.cdc.gov/travel/yellowbook/2024/infections-diseases/melioidosis#agent

Luque-Fernandez, M.A. 2018. *Introduction to Spatial Epidemiology Analyses and Methods.* Instituto de Investigacion Biosanitaria (ibs.GRANADA) chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://scholar.harvard.edu/files/malf/files/spepi.pdf

Maryam Ahmed Alramah, Nor Azah Samat & Zulkifley Mohamed. 2019. Mapping lung cancer disease in Libya using Standardized Morbidity Ratio, BYM Model and Mixture Model, 2006 to 2011: Bayesian epidemiological study. *Sains Malaysiana* 48(1): 217-225 http://dx.doi.org/10.17576/jsm-2019-4801-25

Melikechi, O., Young, A.L., Tang, T., Bowman, T., Dunson, D. & Johndrow, J. 2022. Limits of epidemic prediction using SIR models. *J. Math. Biol.* 85(4): 36. doi: 10.1007/s00285-022-01804-5

Melioidosis.info. 2024. *Cases Information*. https://www.melioidosis.info/info.aspx?pageID=107

Moraga, P. 2019. *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*. New York: Chapman & Hall. https://www.paulamoraga.com/book-geospatial/sec-arealdatatheory.html

Nathan, S., Chieng, S., Kingsley, P.V., Mohan, A., Podin, Y., Ooi, M.H., Mariappan, V., Vellasamy, K.M., Vadivelu, J., Daim, S. & How, S.H. 2018. Melioidosis in Malaysia: Incidence, clinical challenges, and advances in understanding pathogenesis. *Trop. Med. Infect. Dis.* 3(1): 25. doi: 10.3390/tropicalmed3010025

Nicholls, A. 2020. *The Standardized mortality ratio and how to calculate it*. https://s4be.cochrane.org/blog/2020/08/26/the-standardised-mortality-ratio-and-how-to-calculate-it/

Phillips, E.D. & Garcia, E.C. 2024. *Burkholderia pseudomallei*. *Trends in Microbiology* 32(1): 105-106. https://www.cell.com/trends/microbiology/fulltext/S0966-842X(23)00205-6

Puthucheary, S.D. 2009. Melioidosis in Malaysia. *Med. J. Malaysia* 64(4): 266-274.

Shaktawat, Y.S. 2020. *What is ArcGIS?* https://doc.arcgis.com/en/arcgis-online/get-started/what-is-agol.htm

Sia, T.L.L., Mohan, A., Ooi, M.H., Chien, S.L., Tan, L.S., Goh, C., Pang, D.C.L., Currie, B.J., Wong, J.S. & Podin, Y. 2021. Epidemiological and clinical characteristics of melioidosis caused by gentamicin-susceptible *Burkholderia pseudomallei* in Sarawak, Malaysia. *Open Forum Infect Dis.* 8(10): ofab460. doi: 10.1093/ofid/ofab460. Erratum in *Open Forum Infect Dis*. 9(2): ofab653. doi: 10.1093/ofid/ofab653

Trejo, I. & Hengartner, N.W. 2022. A modified susceptible-infected-recovered model for observed under-reported incidence data. *PLoS ONE* 17(2): e0263047. doi: 10.1371/journal.pone.0263047

University of California Santa Cruz (UC Santa Cruz). (n.d.). *DIC: Deviance Information Criterion*. https://ams132-winter17-01.courses.soe.ucsc.edu/home.html

*Corresponding author; email: nazrina@uum.edu.my